

Linearized Jastrow-style fluctuations on spin-projected Hartree-Fock

Thomas M. Henderson and Gustavo E. Scuseria

Department of Chemistry and Department of Physics and Astronomy, Rice University, Houston, Texas 77251-1892, USA

(Received 9 October 2013; accepted 2 December 2013; published online 19 December 2013)

The accurate and efficient description of strong electronic correlations remains an important objective in electronic structure theory. Projected Hartree-Fock theory, where symmetries of the Hamiltonian are deliberately broken and projectively restored, all with a mean-field computational scaling, shows considerable promise in this regard. However, the method is neither size extensive nor size consistent; in other words, the correlation energy per particle beyond broken-symmetry mean field vanishes in the thermodynamic limit, and the dissociation limit of a molecule is not the sum of the fragment energies. These two problems are closely related. Recently, Neuscamman [Phys. Rev. Lett. **109**, 203001 (2012)] has proposed a method to cure the lack of size consistency in the context of the antisymmetrized geminal power wave function (equivalent to number-projected Hartree-Fock-Bogoliubov) by using a Jastrow-type correlator in Hilbert space. Here, we apply the basic idea in the context of projected Hartree-Fock theory, linearizing the correlator for computational simplicity but extending it to include spin fluctuations. Results are presented for the Hubbard Hamiltonian and for some simple molecular systems. © 2013 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4848075>]

I. INTRODUCTION

While the quantum chemistry community has effectively solved the electronic structure problem for weakly correlated electrons, the same cannot be said when electronic correlations are strong and the familiar mean-field picture breaks down. There are, of course, a wide variety of methods which can be applied to strongly correlated systems, including complete active space self-consistent field (CAS-SCF),^{1,2} density matrix renormalization group (DMRG),³ and several others. All have limitations and involve choices that are not yet fully black box.

Recently, we and others have demonstrated⁴ an efficient approach to projected Hartree-Fock (PHF).⁵⁻⁸ In PHF, the ground state wave function $|\Psi\rangle$ is taken to be the result of projecting a broken symmetry determinant $|\Phi\rangle$ such that $|\Psi\rangle$ is a symmetry eigenfunction:

$$|\Psi\rangle = \hat{P}|\Phi\rangle, \quad (1)$$

where $\hat{P} = \hat{P}^\dagger = \hat{P}^2$ is a symmetry projection operator.⁹ The energy is taken to be the expectation value of the Hamiltonian with respect to $|\Psi\rangle$,

$$E_\Psi = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\langle\Phi|\hat{H}\hat{P}|\Phi\rangle}{\langle\Phi|\hat{P}|\Phi\rangle}, \quad (2)$$

and we determine $|\Phi\rangle$ by variationally minimizing this projected energy in a variation after projection scheme.

The PHF wave function generally describes strong correlations in finite systems reasonably well, and has provided highly accurate results across a variety of strongly correlated systems.^{8,10-13} Only a few choices are required: one needs only to select the symmetries which $|\Phi\rangle$ is allowed to break and the symmetries which the projector restores, along with the quantum numbers associated with these symmetries. Fur-

thermore, symmetries that break spontaneously (such as spin) seem to be most important.

However, PHF is not without drawbacks. There are three of particular importance. First, it is not really suitable to the description of dynamical correlations. Second, it is not size extensive:¹⁴ in the thermodynamic limit, the PHF correlation energy is a constant so that the correlation energy per electron vanishes and the energy per particle is that of broken-symmetry Hartree-Fock. Finally, PHF is not size consistent, where we understand size consistency as energy additivity for non-interacting systems; thus, a molecule described by PHF does not dissociate into a set of PHF fragments. These problems are not unconnected—making PHF size consistent would require adding additional correlations to the method and would of course be closely associated with size extensivity as well.

The failure of PHF to be size consistent is easily understood. At the dissociation limit, describing each fragment at the PHF level would require each fragment to have its own projection operator. Thus, in the dissociation of CO₂ to CO and the oxygen atom, we should really insist that each individual fragment be a spin eigenfunction. Instead, PHF insists that the dissociated supermolecule be a spin eigenfunction but does not project each fragment separately. The result is that the PHF dissociation limit is too high because of symmetry fluctuations between non-interacting fragments which cannot be suppressed by a global projection operator. In a sense, this is like the failure of truncated configuration interaction, in which, for example, the double excitation operator cannot create simultaneous double excitations on each fragment, which is a quadruple excitation that would be required for size consistency.

While it is clear that using local projection operators would solve the problem of size consistency at dissociation, the question then becomes how one would move smoothly

from local projection operators at dissociation to a global projection operator at equilibrium. Thus, that we know the origin of the lack of size consistency in PHF is not enough. We must also find some way of overcoming the problem. In the case of configuration interaction, this was provided by exponentiating the correlation operator to give coupled cluster theory, but it is not entirely clear how to do this with projection operators or even whether it is possible at all.

Such a way has apparently been provided by Neuscamman^{15–17} in the case of the antisymmetrized geminal power (AGP) wave function, which is identical to number-projected Hartree-Fock-Bogoliubov.^{6,7,18} From the symmetry projection point of view, the AGP is not size consistent because each individual dissociation fragment is not a number eigenfunction while the AGP wave function is. This can be remedied by noting that

$$\hat{Q} = e^{-\alpha(\hat{N}-N)^2} \quad (3)$$

is a projector as α goes to infinity, in that it eliminates all components of a wave function except those that are eigenfunctions of the operator \hat{N} with eigenvalue N . One could imagine a product of such operators acting on separate Hilbert spaces, via

$$\hat{Q} = \prod_i e^{-\alpha_i(\hat{N}_i - N_i)^2} = e^{-\sum_i \alpha_i(\hat{N}_i - N_i)^2} \quad (4)$$

where \hat{N}_i acts only on states in the i th space. This operator \hat{Q} would then take the form of a product of local projectors at dissociation but could act as a global projector near equilibrium since $\sum \hat{N}_i = \hat{N}$. We could then write a wave function ansatz of the form

$$|\chi\rangle = \hat{Q}|\text{AGP}\rangle. \quad (5)$$

In practice, it is more convenient to replace \hat{Q} with the related operator

$$e^{\hat{J}} = e^{\sum J_{PQ}\hat{N}_P\hat{N}_Q + \sum J_P\hat{N}_P}, \quad (6)$$

where \hat{N}_P and \hat{N}_Q are number operators for orbitals P and Q . Up to an irrelevant multiplicative constant, this operator contains \hat{Q} in the limit of appropriately chosen coefficients J_{PQ} and J_P , and therefore can yield a size consistent result. It in fact does so when the coefficients are chosen variationally, since size inconsistency necessarily raises the ground state energy. This operator is akin to a Hilbert-space analogue of a Jastrow function, for which reason the technique has been dubbed a Hilbert space Jastrow correlator.

Inspired by this Jastrow-type ansatz, we pursue a similar course with PHF as a reference rather than AGP. The chief difficulty lies in handling the exponential operator, which Neuscamman dealt with in a variational approach by a Monte Carlo scheme. Here, we investigate the drastic approximation of replacing the exponential with a linear operator: $e^{\hat{J}} \rightarrow 1 + \hat{J}$. We can then optimize the coefficients in \hat{J} with a straightforward matrix diagonalization. We would expect to lose size consistency, for which the exponential ansatz is crucial, but we hope to see significant improvements upon PHF nevertheless.

The remainder of this paper is organized as follows. In Sec. II we first review the basics of projected Hartree-Fock

theory, then discuss the correlator at some length. Section III discusses results in the Hubbard Hamiltonian and in a few molecular systems, and we conclude in Sec. IV.

II. THEORY

While there are many symmetries we could consider in PHF, we will confine ourselves to the simplest PHF model: spin-projected unrestricted Hartree-Fock (SUHF), where the mean-field determinant $|\Phi\rangle$ is an eigenfunction of \hat{S}_z and we project onto eigenfunctions of \hat{S}^2 . The manner in which this projection is carried out is not entirely trivial, so we will begin by outlining SUHF. That done, we will turn to discuss the form of the Jastrow-style correlator \hat{J} which we will consider and then sketch our working equations. They will require matrix elements which are quite complicated, and we will accordingly not present the matrix elements in detail. The computational cost, however, scales as $\mathcal{O}(N^5 N_{\text{grid}})$, where N is the number of basis functions of the system and N_{grid} is the number of points needed for the numerical quadrature outlined below. See the Appendix for details.

A. Spin-projected UHF

The basic idea in SUHF is to write a UHF-style single determinant $|0\rangle$ which is then spin projected to form an eigenfunction of \hat{S}^2 . The way in which this is actually accomplished is somewhat indirect. It is not our intention to provide a detailed overview of PHF, which can be found in Ref. 4.

In ordinary PHF, one could write the projection operator in the form

$$\hat{P} = \int d\theta e^{i\theta(\hat{S}^2 - \langle\hat{S}^2\rangle)}, \quad (7)$$

where $\langle\hat{S}^2\rangle$ is the desired expectation value. However, because \hat{S}^2 is a two-particle operator, the action of this projection operator on a single determinant would be to produce a linear combination of up to N -tuply excited states, where N is the number of electrons in the determinant. This approach is therefore to be avoided at all costs.

Instead, we take advantage of the SU(2) structure of spin and make the projected wave function invariant under all rotations of the spin quantization axis. In doing so, we achieve a spin eigenfunction. This works through writing a projection-like operator

$$\hat{P} = \int d\Omega W(\Omega) \hat{R}(\Omega), \quad (8)$$

where Ω stands for a trio of angles (α, β, γ) , the rotation operator $\hat{R}(\Omega)$ is

$$\hat{R}(\Omega) = e^{i\alpha\hat{S}_z} e^{i\beta\hat{S}_y} e^{i\gamma\hat{S}_z}, \quad (9)$$

and the integration weights $W(\Omega)$ are the matrix elements of this rotation operator between spin eigenstates.

This projection-like operator can be simplified considerably in the case of a UHF reference determinant. In this case, we can reduce the evaluation of the energy to a single

integration by using

$$\hat{P}' = \int d\beta \tilde{W}(\beta) e^{i\beta \hat{S}_y}. \quad (10)$$

Here, we have used the facts that the reference determinant is an eigenfunction of $\exp(i\gamma \hat{S}_z)$ and that $\exp(i\alpha \hat{S}_z)$ commutes with the Hamiltonian to do the integrations with respect to α and γ analytically; $\tilde{W}(\beta)$ is what remains of $W(\Omega)$ after these integrations are done. It is important to remember, however, that the projected wave function has the more complicated structure associated with the actual projector of Eq. (8). This is because when we evaluate the expectation value of a general operator, we have

$$\begin{aligned} \langle 0 | \hat{P}^\dagger \hat{O} \hat{P} | 0 \rangle &= \int d\Omega d\Omega' W(\Omega) W^*(\Omega') \\ &\times \langle 0 | \hat{R}^\dagger(\Omega') \hat{O} \hat{R}(\Omega) | 0 \rangle. \end{aligned} \quad (11)$$

While the integrations over γ and γ' can be done analytically, the remaining four integrations still must be carried out. Only when the operator \hat{O} commutes with the projector \hat{P} is it possible to evaluate the expectation value using a single integration.

To evaluate general expectation values, it is helpful to consider the special case when \hat{O} is just a string of creation and annihilation operators, which we will refer to as a density matrix operator $\hat{\Gamma}_k$. By this, we mean that a k -particle reduced density matrix is schematically

$$\Gamma_k = \frac{1}{k!} \frac{\langle 0 | \hat{P}^\dagger \hat{\Gamma}_k \hat{P} | 0 \rangle}{\langle 0 | \hat{P} | 0 \rangle}. \quad (12)$$

We can then evaluate the expectation value of a k -particle operator \hat{O}_k as simply

$$\langle \hat{O}_k \rangle = \frac{1}{k!} \text{Tr}(\Gamma_k \mathbf{O}_k), \quad (13)$$

where \mathbf{O}_k contains antisymmetrized matrix elements of the operator \hat{O}_k .

The density matrix Γ_k can be written as an integration over the projection grid of a transition density matrix γ_k :

$$\Gamma_k = \int W(\Omega) W^*(\Omega') \gamma_k(\Omega, \Omega'), \quad (14)$$

where the angle Ω stands for (α, β) and the transition density matrix is just

$$\gamma_k(\Omega, \Omega') = \langle 0 | e^{-i\beta' \hat{S}_y} e^{-i\alpha' \hat{S}_z} \hat{\Gamma}_k e^{i\alpha \hat{S}_z} e^{i\beta \hat{S}_y} | 0 \rangle. \quad (15)$$

Because the PHF wave function is a sum of single determinants created by the action of sequential orbital rotations, the k -particle transition density matrix γ_k factorizes nicely, just as it does in Hartree-Fock:¹⁸

$$\gamma_k = \frac{1}{k!} \underbrace{\gamma_1 \wedge \gamma_1 \wedge \dots}_{k \text{ times}} \quad (16)$$

where the notation is meant to imply that we have the antisymmetrized product of k copies of γ_1 . In turn, γ_1 can be evaluated as we have already outlined. As long as we are interested only in operators which commute with the projectors, we can consider only the symmetry-adapted part of the density matrix, which permits us to evaluate γ_1 and therefore Γ_k

using single integration with the operator \hat{P}' of Eq. (10). This is the strategy we will take throughout this work. Note that by the symmetry-adapted part of the density matrix, we mean the portion which transforms as the totally symmetric irreducible representation.

We should say a few words about the numerical integration required. The weights are ultimately just trigonometric functions – sines and cosines and products thereof – which must be integrated over the unit circle. We thus use simple Gauss-Legendre quadrature. We find that 16 points suffice for any of the systems that we study in this work, though we note that for larger systems, somewhat more points may be needed.

B. The Jastrow-type correlator

The correlated wave function we wish to consider is

$$|\Psi\rangle = (1 + \hat{J})\hat{P}|0\rangle = \hat{J}\hat{P}|0\rangle \quad (17)$$

where we will write

$$\hat{J} = \sum \mathcal{J}_\mu \hat{O}_\mu \quad (18)$$

and will defer consideration of the form of the operators \hat{O}_μ until later. The coefficients \mathcal{J}_μ are to be determined variationally, which naturally leads to a generalized eigenvalue problem

$$H_{\mu\nu} \mathcal{J}_\nu = S_{\mu\nu} \mathcal{J}_\nu E, \quad (19)$$

where

$$H_{\mu\nu} = \langle 0 | \hat{P}^\dagger \hat{O}_\mu^\dagger \hat{H} \hat{O}_\nu \hat{P} | 0 \rangle, \quad (20a)$$

$$S_{\mu\nu} = \langle 0 | \hat{P}^\dagger \hat{O}_\mu^\dagger \hat{O}_\nu \hat{P} | 0 \rangle. \quad (20b)$$

We can evaluate these expectation values with the aid of the SUHF density matrix as outlined above.

Now, the reference wave function $\hat{P}|0\rangle$ is a symmetry eigenfunction, and because we are diagonalizing the Hamiltonian in the basis of states $\{\hat{P}|0\rangle, \hat{O}_\mu \hat{P}|0\rangle\}$ we expect the correlated wave function $|\Psi\rangle$ to be a symmetry eigenfunction as well if the operators \hat{O}_μ permit it. It is then clear that we need to consider only symmetry-adapted operators \hat{O}_μ (*i.e.*, operators which are totally symmetric). Indeed, if we do not symmetry adapt \hat{O}_μ we will find relations amongst the coefficients \mathcal{J}_μ which return the symmetry-adapted result (again, assuming that the symmetry-adapted result is possible to construct from the chosen operators). That \hat{O}_μ must be symmetry adapted is quite useful in our case, as it allows us to evaluate the necessary matrix elements with the single integration we have discussed earlier. To put it another way, if the operators \hat{O}_μ transform as the totally symmetric irreducible representation, they commute with the projection operators \hat{P} and we can then write, for example,

$$H_{\mu\nu} = \langle 0 | \hat{O}_\mu^\dagger \hat{H} \hat{O}_\nu \hat{P} | 0 \rangle. \quad (21)$$

The need to symmetry adapt \hat{O}_μ constrains our flexibility. This flexibility is further constrained by the notion that the operators should in some sense be orbital fluctuations, as

in Neuscamman's work, and still further by the desire to restrict ourselves to one- and two-body operators. In fact, the only one-body spin singlet orbital fluctuation operators we can construct are

$$\hat{N}_P = a_{P\uparrow}^\dagger a_{P\uparrow} + a_{P\downarrow}^\dagger a_{P\downarrow}, \quad (22)$$

where P denotes a spatial orbital of the system. We can add the two-body operators $\hat{N}_P \hat{N}_Q$ which together with \hat{N}_P yield Neuscamman's original correlator and act to damp local charge fluctuations as we have discussed above. The only other two-body spin singlet operators we can construct are

$$\sum_k \hat{S}_P^k \hat{S}_Q^k = \sum_k \sum_{\lambda, \delta, \mu, \nu} \sigma_{\lambda\delta}^k \sigma_{\mu\nu}^k a_{P\lambda}^\dagger a_{P\delta} a_{Q\mu}^\dagger a_{Q\nu}, \quad (23)$$

where the indices λ, δ, μ , and ν indicate spin directions and σ^k is the k th Pauli matrix. These operators act to damp local spin fluctuations. Any other one- or two-body operators one could form either break spin symmetry or do not take the Jastrow-like form in which we are interested (that is, they do not take the form of orbital fluctuations). Explicitly, we have

$$1 + \hat{J} = 1 + \sum_{P \geq Q} J_{PQ} \hat{N}_P + \sum_{P \geq Q} J_{PQ} \hat{N}_P \hat{N}_Q \quad (24a)$$

$$+ \sum_{P \geq Q} K_{PQ} \hat{S}_P \cdot \hat{S}_Q$$

$$= 1 + \sum_{P \geq Q} J_{PQ} \hat{N}_P + \sum_{P > Q} J_{PQ} \hat{N}_P \hat{N}_Q$$

$$+ \sum_{P > Q} K_{PQ} \hat{S}_P \cdot \hat{S}_Q + \sum_P K_P \hat{N}_{P\uparrow} \hat{N}_{P\downarrow}. \quad (24b)$$

Here, we have used the fact that \hat{N}_P and \hat{N}_Q commute for $P \neq Q$ and similarly for \hat{S}_P and \hat{S}_Q , so that we need consider only $P \geq Q$ in the quadratic operators. For $P = Q$, we have that \hat{N}_P , \hat{N}_P^2 , and $\hat{S}_P \cdot \hat{S}_P$ are not linearly independent and the linearly independent operators can be chosen as \hat{N}_P and $\hat{N}_{P\uparrow} \hat{N}_{P\downarrow}$.

It is important to note that the correlator is not invariant to rotations amongst the orbitals in the operators. That is, \hat{N}_P measures the number of electrons in the spatial orbital ϕ_P , which then forces us to specify the orbital. If, for instance, the orbital ϕ_P is a natural orbital (*i.e.*, an eigenfunction of \hat{N}_P) then the correlator has no effect on a single determinant reference unless we include the spin fluctuations. This should be contrasted to the more familiar correlation operators in coupled cluster theory or configuration interaction, where there is an invariance to occupied-occupied or virtual-virtual orbital mixing. Physically, it would appear to be sensible that these orbitals should be spatially localized in nature, so that as we dissociate a molecule the orbitals localize on one of the fragments; after all, the idea is that the correlator should become a projector were it exponentiated. On the other hand, variational considerations suggest that in some way the orbitals should be as unlike the natural orbitals as possible. Ideally, one would optimize the choice of these orbitals, but this variational problem would be exceedingly complicated and we will not consider it further here. We will note the identity of the orbitals being used to define the correlator as we discuss our results.

The second important note is that the correlator does not include general purpose single excitations. Therefore, we expect to see rather different results depending on the choice of projected reference wave function or, to put it differently, depending on the choice of the broken symmetry mean-field reference determinant $|0\rangle$. Unless otherwise stated, we will use the reference determinant of SUHF so that $\hat{P}|0\rangle = |\text{SUHF}\rangle$.

Finally, we note that because the correlator involves up to two-particle operators \hat{O}_{PQ} with two spatial orbital indices, evaluation of Hamiltonian matrix elements will require up to six-particle density matrices. The apparent scaling is then $\mathcal{O}(N^8 N_{\text{grid}})$, where we get $\mathcal{O}(N^2)$ each from \hat{O}_{PQ} and \hat{O}_{RS} , and a further $\mathcal{O}(N^4)$ from the Hamiltonian. However, by using the decomposition of the density matrix into the antisymmetrized product of one-particle transition density matrices, we find that computational scaling is just $\mathcal{O}(N^5 N_{\text{grid}})$, as shown in the Appendix.

III. RESULTS

In this section, we will discuss illustrative results of this Jastrow-type correlator formalism applied to SUHF wave functions. We will begin with the Hubbard Hamiltonian before moving on to molecular systems. All calculations were done with an in-house code and are on spin singlets for simplicity of the SUHF calculation and density matrix. We will refer to SUHF with the Jastrow-like correlator added as JSUHF; if the correlator includes only number fluctuations, as originally proposed, we will denote it by JSUHF_N.

A. The Hubbard Hamiltonian

The Hubbard Hamiltonian¹⁹ is a simple but rich lattice model which describes the essential competition between itinerancy and localization favored by the kinetic energy and electron-electron repulsion, respectively. We have

$$\hat{H} = -t \sum_{\langle IJ \rangle} \sum_{\sigma} a_{I\sigma}^\dagger a_{J\sigma} + U \sum_I \hat{N}_{I\uparrow} \hat{N}_{I\downarrow}, \quad (25)$$

where in the first term the summation over σ runs over spins and the summation over sites I and J runs only over nearest-neighbor sites. We will be interested here in one-dimensional lattices with N sites, typically with periodic boundary conditions so that site $N + 1$ is equivalent to site 1. The orbitals defining, for example, \hat{N}_I in the correlator will always be these localized sites. We will take $t = 1$ or, in other words, will measure the energy E and the constant U in units of t .

We begin with two-electron systems in Table I. For two sites, SUHF is already exact, but for more sites, SUHF deteriorates until, for a sufficiently large number of sites, spin projection adds only a negligible fraction of the correlation energy beyond broken-symmetry mean-field. Adding the correlator, however, gives the exact result, with or without spin fluctuations. This should not be too surprising. For N spatial orbitals (here, N sites), the number of two-electron singlet states is $1/2 N(N + 1)$, which is the same as the number of parameters in $1 + \sum_P J_P \hat{N}_P + \sum_{P > Q} J_{PQ} \hat{N}_P \hat{N}_Q$ and thus less

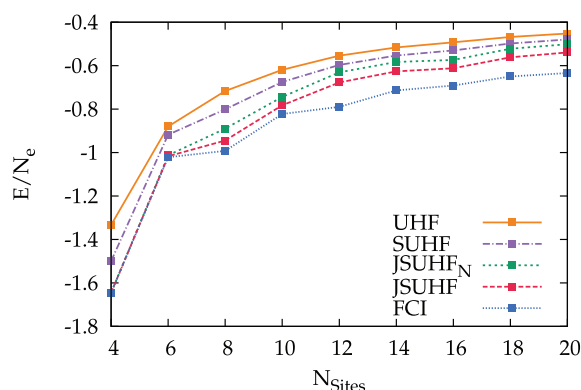
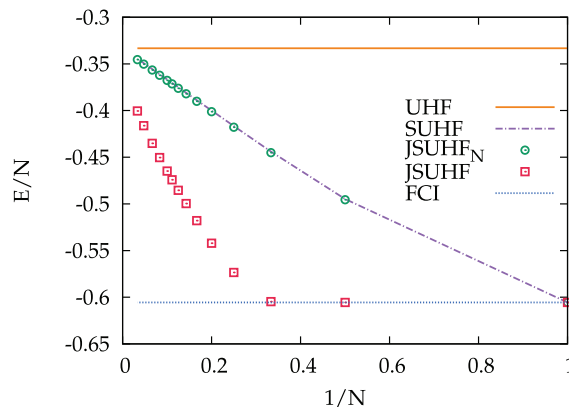
TABLE I. Energy for a two-electron Hubbard ring with N sites. Singlet state, $U = 6$.

N	UHF	SUHF	JSUHF	Exact
2	-1.333333	-2.000000	-2.000000	-2.000000
4	-2.666667	-3.000000	-3.291503	-3.291503
6	-3.232702	-3.385441	-3.631298	-3.631298
8	-3.508587	-3.583282	-3.773325	-3.773325
10	-3.659423	-3.698742	-3.846456	-3.846456
12	-3.750050	-3.771857	-3.889101	-3.889101

than the number of parameters in the correlator as a whole. Of course we cannot guarantee that we must therefore get the exact result because, depending on the choice of orbitals involved in constructing the correlator, we may or may not obtain $1/2 N(N+1)$ linearly independent vectors, but in practice we do not expect to encounter this difficulty (and have never done so).

The two-electron results serve mainly to validate the code, but do not provide much insight. After all, SUHF is already fairly accurate for these two-electron systems—while the SUHF correlation energy becomes small, so too does the total correlation energy. It is more useful to examine the two-hole states (that is, the lowest energy singlet state for $N-2$ electrons in an N -site Hubbard system). The presence of these holes creates a strong correlation less readily captured by spin projected methods. Results are presented in Fig. 1. Note that now the correlation energy is a much larger fraction of the total energy, and as N becomes large, the fraction of the correlation energy recovered by SUHF becomes small. Adding the correlator helps significantly, particularly when we include both number and spin fluctuations. Indeed, at $N=20$ the SUHF recovers only 15% of the correlation energy while the JSUHF recovers nearly half of it.

All of this shows that adding the correlator captures some of the additional correlations on top of SUHF – correlations we will refer to as “residual correlations” – but does not address the issues of size consistency or size extensivity. We do not expect complete success in this regard, as we have noted before, simply because the correlator is not exponential. However, the presence of the spin fluctuations $\hat{S}_P \cdot \hat{S}_Q$ provides

FIG. 1. Energy per electron in an N -site Hubbard ring with two holes (*i.e.*, $N-2$ electrons). Singlet state, $U = 6$.FIG. 2. Energy per fragment for an N -fragment Hubbard dimer model at half-filling, plotted against $1/N$. Singlet state, $U = 6$.

surprisingly large improvements for these properties as well, even with this simple linear ansatz.

To examine size consistency, we consider the energy per fragment for a Hubbard dimer model (Fig. 2). In the dimer model, each site is connected to only one other site; we will refer to this pair of sites as a fragment. The various fragments are completely noninteracting. This being the case, a size consistent method would predict that the energy per fragment for N fragments would be N times the energy of a single fragment or, putting it another way, a plot of the energy per fragment against the number of fragments would give a straight line. This is precisely what one sees for UHF and for the exact result. In contrast, the SUHF energy per fragment goes from the exact result for one fragment to the UHF result for a large number of fragments. SUHF, in other words, is not size consistent. Nor is JSUHF_N, which is virtually indistinguishable from SUHF. Adding the spin fluctuations means that JSUHF is energetically exact for one and two fragments and very close to exact for three fragments before eventually deteriorating. Both SUHF and JSUHF_N approach UHF as $1/N$, indicating that the total correlation energy in SUHF and JSUHF_N is a constant (as we expect for SUHF). The full JSUHF appears to approach UHF with more of a \sqrt{N} -like behavior, as we expect for truncated configuration interaction.

To examine size extensivity, we look at the energy per electron for a Hubbard ring at half filling (where the number of electrons is equal to the number of sites). Results are shown in Fig. 3. As one can see, the UHF and the exact result very quickly saturate. As with the dimer model, the SUHF approaches UHF as $1/N$, demonstrating that in the thermodynamic limit, SUHF is below UHF by a constant. Without spin fluctuations, the JSUHF_N differs only slightly from SUHF. Adding spin fluctuations again means that JSUHF approaches UHF with a $1/\sqrt{N}$ -type behavior. Thus, while the linear correlator we have used does not restore size extensivity, it offers substantial improvements upon SUHF in this respect, provided that we include the spin fluctuations.

Note that Fig. 3 adds curves marked “LMSUHF” and “JLMSUHF.” These incorporate the effects of lattice momentum projection.¹⁰ In Hubbard rings, the lattice momentum is a symmetry (*i.e.*, states can be labeled by a lattice vector \mathbf{k}), and this symmetry can be projected. Earlier research has indi-

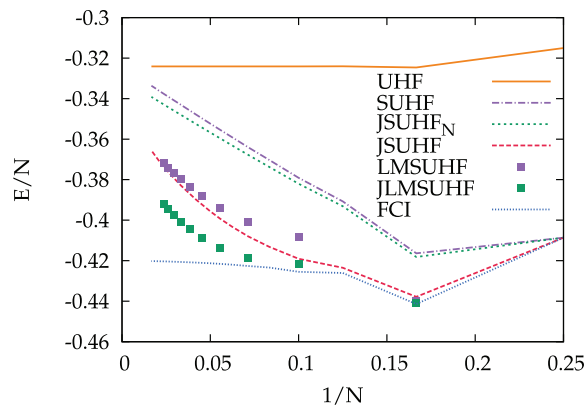


FIG. 3. Energy per electron in an N -site Hubbard ring at half filling, plotted against $1/N$. Singlet state, $U = 6$.

cated that lattice momentum projection is quite important for the Hubbard Hamiltonian.^{10,11} We should point out that the projection increases the scaling of the projected Hartree-Fock with system size, simply because the number of grid points in the lattice momentum projection quadrature must be the number of lattice sites. The requirement that the Jastrow-style correlator must have $\mathbf{k} = \mathbf{0}$ means that we can write it as

$$1 + \hat{J} = 1 + \sum_P \sum_{q>0} (J_q \hat{N}_P \hat{N}_{P+q} + K_q \hat{S}_P \cdot \hat{S}_{P+q}), \quad (26)$$

where the sum over q runs over site separation vectors, so that $q = 1$ corresponds to nearest neighbors, $q = 2$ to next-nearest neighbors, and so forth. Clearly, there are fewer variational parameters available for the correlator, which is a general phenomenon: as the number of symmetries being projected is increased, the number of constraints on the correlator likewise increases, and the number of variational parameters decreases. Put another way, symmetry adapting the correlator means we diagonalize a smaller block of the Hamiltonian matrix. Results for lattice projection are presented in Fig. 3 as LMSUHF (for the spin and lattice-momentum projected Hartree-Fock) and JLMSUHF (for the addition of the correlator). In this case, we project only onto states with momentum $\mathbf{k} = \mathbf{0}$, which are ground states for $4n + 2$ sites at half-filling. We note that the LMSUHF is of accuracy comparable to that of JSUHF, while adding the correlator further decreases the

errors. On the other hand, the method is still not size extensive – the linear correlator simply lacks the flexibility to remedy this defect.

We have said that the reference determinant $|0\rangle$ will have potentially large effects on the JSUHF-type wave function. To assess these effects, we consider the results of using several different reference determinants. While SUHF refers to spin-projected UHF with variation after projection, we will use PUHF to denote spin-projected UHF with projection after variation. In other words, the reference determinant of PUHF is the UHF determinant. We use JPUHF to refer to PUHF with the Jastrow-type correlator added. We also consider results obtained by optimizing the identity of $|0\rangle$ by Jacobi rotations starting from SUHF. This method we refer to as JOpt.

Results for Hubbard rings are shown in Table II. We see that the choice of the reference determinant can have a pronounced effect on the results in the absence of spin fluctuations in the correlator, accounting for up to $\sim 15\%$ of the correlation energy with respect to UHF. Interestingly, adding the spin fluctuations seems to wash most of this dependence out. We do not at present have a good explanation for why this should be.

Most of the results presented here have been at $U = 6$, in the intermediate coupling regime, which is particularly difficult to capture with mean-field-based methods. We have observed qualitatively similar results for both smaller and larger values of U . Rather than belabor the Hubbard Hamiltonian further, however, we turn to molecular applications.

B. Molecular results

Our studies in molecular systems face one difficulty not present in the Hubbard Hamiltonian. While for the latter, the site basis seems a natural choice for a basis in which to express the correlator, we do not have such a luxury in molecules. The atomic orbital basis might seem to be ideal, but because it is not orthonormal, it complicates the equations significantly. We choose to work with the symmetrically orthonormalized atomic orbitals, which write

$$|\phi_p\rangle = X_{\mu p} |\chi_\mu\rangle, \quad (27)$$

TABLE II. Effects of orbital choice for an N -site Hubbard ring. Singlet state, $U = 6$.

N	UHF	PUHF	SUHF	JPUHF _N	JSUHF _N	JOpt _N	JPUHF	JSUHF	JOpt	Exact
Half filling										
2	-1.333333	-1.846154	-2.000000	-2.000000	-2.000000	-2.000000	-2.000000	-2.000000	-2.000000	-2.000000
4	-1.260142	-1.606487	-1.634544	-1.634603	-1.634603	-1.634603	-1.634603	-1.634603	-1.634603	-1.634603
6	-1.947605	-2.488520	-2.498192	-2.514828	-2.508960	-2.569752	-2.621954	-2.626947	-2.631818	-2.648518
8	-2.592485	-3.119006	-3.125755	-3.156842	-3.144973	-3.219563	-3.388033	-3.377877	-3.389114	-3.408772
10	-3.240981	-3.785380	-3.790757	-3.833577	-3.820339	-3.923631	-4.189335	-4.190726	-4.194153	-4.254559
Two hole										
4	-2.666667	-2.923077	-3.000000	-3.291503	-3.291503	-3.291503	-3.291503	-3.291503	-3.291503	-3.291503
6	-3.520549	-3.603949	-3.672172	-4.058705	-4.050333	-4.059892	-4.069485	-4.068243	-4.075478	-4.084766
8	-4.299689	-4.748317	-4.806564	-5.419688	-5.338032	-5.614081	-5.596443	-5.668536	-5.730032	-5.954448
10	-4.952662	-5.374353	-5.402281	-6.033192	-5.952697	-6.179947	-6.264620	-6.259378	-6.312982	-6.578259

TABLE III. Correlation energies (mH) in OH^- and H_2O at their equilibrium geometries using the cc-pVDZ basis.²⁰ For JSUHF and JSUHF_N, we show both symmetrically orthogonalized AOs (SAO) and canonically orthogonalized AOs (CAO).

System	SUHF	SAO		CAO	
		JSUHF _N	JSUHF	JSUHF _N	JSUHF
OH^-	40.40	114.08	132.37	73.54	102.96
H_2O	41.95	112.91	127.48	89.05	107.32

where $|\chi_\mu\rangle$ is an atomic orbital. The coefficient matrix \mathbf{X} is determined from

$$\mathbf{X} = \mathbf{U} \mathbf{s}^{-1/2} \mathbf{U}^\dagger, \quad (28)$$

where \mathbf{s} and \mathbf{U} are the eigenvalues and eigenvectors of the overlap matrix of atomic orbitals. Observationally, the symmetrically orthogonalized basis marginally outperforms the canonically orthogonalized basis in which $\mathbf{X} = \mathbf{U} \mathbf{s}^{-1/2}$. For example, Table III shows correlation energies in OH^- and H_2O at equilibrium in the cc-pVDZ basis with the symmetrically and canonically orthogonalized AO bases for the correlator. The effects of the orbitals in which the correlator is expressed are quite pronounced, but while an energetically optimized basis should be used for the correlator, we do not consider that problem here.

Bearing the limitations of correlator orbital choice in mind, we begin by looking at the symmetric double dissociation of H_2O in the cc-pVDZ basis set, as shown in Fig. 4. We use unrestricted CCSD(T) as a reference curve here; this reference is of course imperfect but is a reasonably accurate approximation to the exact result in our basis. It should not be forgotten, however, that JSUHF does not break symmetry, so the more appropriate comparison might be to restricted CCSD(T), which would diverge for large bond lengths. We see that the Jastrow-style correlator is less successful in total energies here than in the case of the Hubbard Hamiltonian, recovering only half of the correlation relative to UCCSD(T). We should note that while neither JSUHF nor SUHF are size consistent, both curves improve significantly upon UHF, and JSUHF makes up much of the difference between the relative energies of CCSD(T) and those of SUHF. Qualitatively similar results are obtained for the dissociation of N_2 , shown in Fig. 5. Again, the Jastrow-like correlator captures roughly half of the correlations beyond SUHF, but offers more significant improvements in relative energies. Curiously, while the spin terms in the correlator have only a small effect on the relative energies in OH^- , they worsen the results in N_2 .

We emphasize one final time that the quality of results obtainable with the Jastrow-type correlator could presumably be improved significantly by optimizing the orbital basis in which the correlator is expressed. We also remind the reader that the computational scaling of the method is $\mathcal{O}(N^5 N_{\text{grid}})$, where N is the size of the underlying basis. Since N_{grid} scales weakly with system size for spin projection, the scaling of JSUHF is certainly less than $\mathcal{O}(N^6)$. We should also point out that the method parallelizes naturally, both in computing ma-

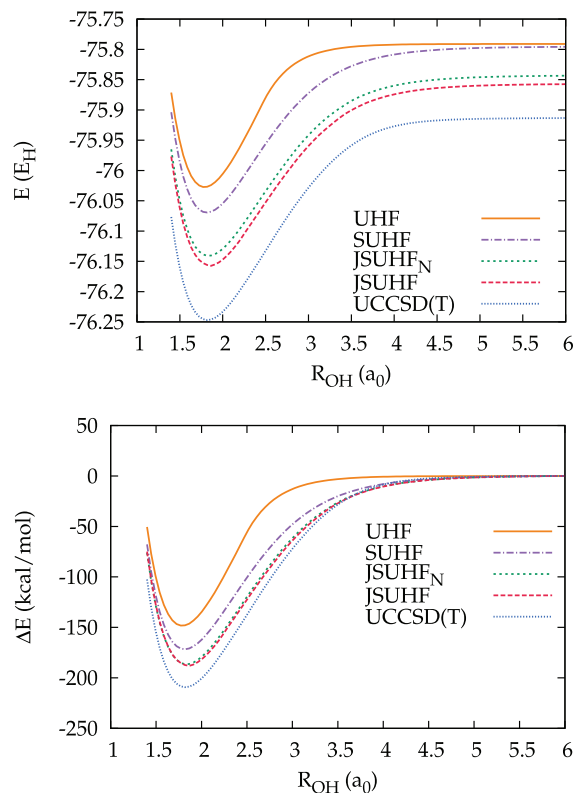


FIG. 4. Symmetric double dissociation of H_2O in the cc-pVDZ basis set. Top panel: Total energies. Bottom panel: Energies relative to the dissociation limit.

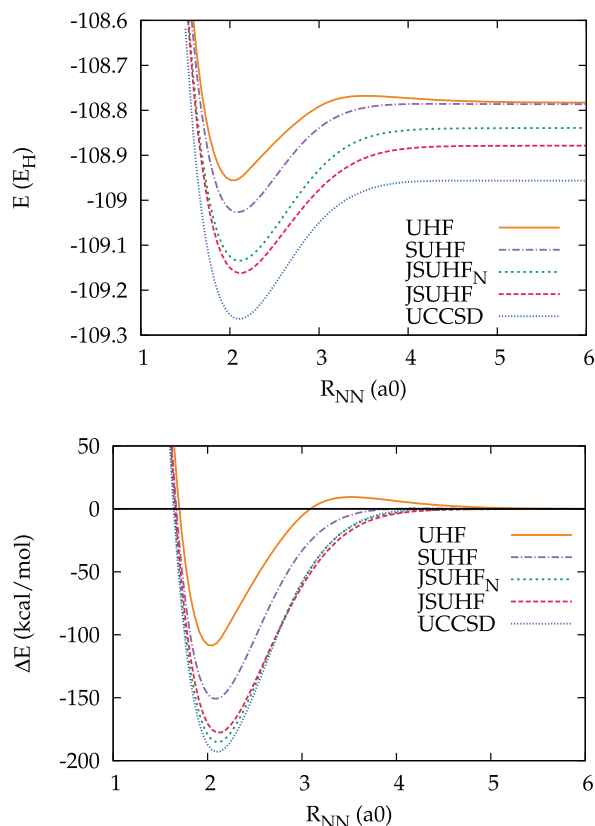


FIG. 5. Dissociation of N_2 in the cc-pVDZ basis set. Top panel: Total energies. Bottom panel: Energies relative to the dissociation limit.

trix elements at individual grid points and across integration grid points.

IV. CONCLUSIONS

One of the great drawbacks of PHF is its size inconsistency and its related size inextensivity. Because it is not size consistent, PHF is not fully reliable in cases where chemical bonds are stretched, which is exactly the regime where strong correlations are most likely to occur. Because it is not size extensive, PHF is best suited to the description of smaller systems. This does not, of course, mean that PHF is not a valuable tool for the description of strongly correlated systems, but it does mean that successors to PHF which address these problems would be significantly more valuable yet.

In the absence of explicit local projection operators, Neuscamman's Hilbert space Jastrow approach seems like a promising way forward. Unfortunately, it requires the variational optimization of coefficients in an exponential operator, and because the exponential operator is not unitary so that $e^{J^\dagger} \neq e^{-J}$, evaluating the expectation value of the wave function is already quite difficult even before we consider the need to differentiate this expectation value. Thus, practical deterministic calculations along these lines must seek an alternative approach.

We have therefore tried the simplest alternative we can imagine—linearizing the exponential. It is clear that in doing so, we will sacrifice the very behavior the Jastrow-type correlator is trying to capture, and that for large enough systems, we will return to PHF (and thus to broken-symmetry mean-field). Nonetheless, we can gain valuable insight with a computationally tractable method which is naturally parallel and scales with system size roughly as does second-order perturbation theory. Indeed, for small systems the problems of size consistency and size extensivity errors are greatly reduced even with this simple linearization. Spin fluctuations seem particularly important in reducing these errors and providing additional residual correlation. Unsurprisingly, excitations rather than orbital fluctuations are required to really get at dynamical correlation in an efficient way, although we note that since the correlator is not written in the molecular orbital basis, single and double excitations lurk inside its form (which is why the results for two-electron systems are exact).

An interesting alternative might be to optimize the coefficients defining the Jastrow-style correlator in the linearized approach we have described and then use these coefficients in a coupled-cluster-like manner, where we evaluate the energy as

$$E = \langle \text{PHF} | e^{-\hat{J}} \hat{H} e^{\hat{J}} | \text{PHF} \rangle. \quad (29)$$

The commutator expansion here would unfortunately not truncate at low-order commutators the way that it does in standard coupled-cluster theory, but one could presumably carry it out until the energy approaches convergence. Such a method might be promising – it would be expected, at least, to ameliorate the size consistency and size extensivity problems of PHF – but its implementation would be quite demanding and auto-

matic program generation would seem to be essential, as neither evaluating the commutator at a given order nor evaluating its expectation value in a computationally tractable framework is straightforward.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy, Office of Basic Energy Sciences, Heavy Element Chemistry program, under Grant No. DEFG02-04ER15523. G.E.S. is a Welch Foundation Chair (C-0036).

APPENDIX: MATRIX ELEMENTS AND COMPUTATIONAL SCALING

We have claimed $\mathcal{O}(N^5 N_{\text{grid}})$ scaling for JSUHF. Here, we wish to demonstrate how this is obtained. We will show results for our most complicated term; other results follow similarly.

Let us begin by switching to a tensor notation which will help simplify our expressions. We define creation operators as

$$a_p^\dagger = c_p \quad (A1)$$

so that the action of c_p on the empty state $|-\rangle$ is to create the state with an electron in spinorbital ϕ_p :

$$c_p |-\rangle = |\phi_p\rangle. \quad (A2)$$

Annihilation operators are then defined with upper indices:

$$a_p = c^p. \quad (A3)$$

In this notation, the Hamiltonian is

$$\hat{H} = h_j^i c_i c^j + \frac{1}{4} v_{kl}^{ij} c_i c_j c^l c^k, \quad (A4)$$

where h_j^i and v_{kl}^{ij} are the one-electron and antisymmetrized two-electron integrals defining \hat{H} . A k -particle density matrix is

$$\Gamma_{p_1 \dots p_k}^{q_1 \dots q_k} = \frac{1}{k!} \langle \Psi | c_{p_1} \dots c_{p_k} c^{q_k} \dots c^{q_1} | \Psi \rangle \quad (A5)$$

and the expectation value of a k -particle operator is

$$\langle \hat{O}_k \rangle = \frac{1}{k!} \Gamma_{p_1 \dots p_k}^{q_1 \dots q_k} O_{q_1 \dots q_k}^{p_1 \dots p_k}, \quad (A6)$$

where $O_{q_1 \dots q_k}^{p_1 \dots p_k}$ are antisymmetrized matrix elements of \hat{O} . In the foregoing equations and indeed throughout this appendix, repeated indices are to be summed.

Now, the two-body operators defining our correlator all take the general form $c_M c^m c_N c^n$ where the notation is meant to indicate that while the *spinorbitals* ϕ_M and ϕ_m may differ, they correspond to the same *spatial* orbital, and likewise for ϕ_N and ϕ_n . Similarly, the one-body operators in the correlator can be written in the general form $c_M c^m$. This notation has the added benefit of emphasizing that there is no sum over m , n , M , or N .

Consider, then, a term like

$$H_{mn,qr} = \langle 0 | c_M c^m c_N c^n \hat{H} c_Q c^q c_R c^r \hat{P} | 0 \rangle. \quad (A7)$$

The highest rank term, which we will denote as $\bar{H}_{mn,qr}$ will come when all creation operators are to the left and all annihilation operators are to the right. We would thus have

$$\begin{aligned}\bar{H}_{mn,qr} &= \langle 0 | c_M c_N c_Q c_R c_i c_j c^r c^q c^n c^m \hat{P} | 0 \rangle h_{ij}^i \\ &+ \frac{1}{4} \langle 0 | c_M c_N c_Q c_R c_i c_j c^l c^k c^r c^q c^n c^m \hat{P} | 0 \rangle v_{kl}^{ij}.\end{aligned}\quad (\text{A8})$$

This we can evaluate as

$$\begin{aligned}\bar{H}_{mn,qr} &= 120 \Gamma_{MNQRi}^{mnqrk} h_k^i + 180 \Gamma_{MNQRij}^{mnqrkl} v_{kl}^{ij} \\ &= \int W(\Omega) (120 \gamma_{MNQRi}^{mnqrk} h_k^i + 180 \gamma_{MNQRij}^{mnqrkl} v_{kl}^{ij}),\end{aligned}\quad (\text{A9})$$

where we have recalled that the density matrices Γ are obtained from the gauge integration of the transition density matrices γ , whose Ω -dependence we have suppressed for brevity.

To proceed, we will use the notion that the transition density matrices γ factor into the antisymmetrized product of transition one-particle density matrices. Taking advantage of this factorization will allow us to evaluate the integrand of Eq. (A9) in $\mathcal{O}(N^5)$ scaling, and therefore all of $\bar{H}_{mn,qr}$ in $\mathcal{O}(N^5 N_{\text{grid}})$ scaling. In what follows, we will therefore work at a single integration point Ω and define

$$\bar{H}_{mn,qr}(\Omega) = 120 \gamma_{MNQRi}^{mnqrk} h_k^i + 180 \gamma_{MNQRij}^{mnqrkl} v_{kl}^{ij}. \quad (\text{A10})$$

Consider, then, the first term, $\gamma_{MNQRi}^{mnqrk} h_k^i$. When we decompose the five-particle density matrix into the antisymmetrized product of one-particle density matrices, the dummy indices i and k could both be on the same one-particle density matrix, or they could be on two different ones. In the former case, we would have a contribution like

$$\delta \bar{H}_{mn,qr}(\Omega) \sim \gamma_{MNQR}^{mnqr} (h_k^i \gamma_i^k), \quad (\text{A11})$$

while in the second we would have something instead like

$$\delta \bar{H}_{mn,qr}(\Omega) \sim \gamma_{NQR}^{nqr} (\gamma_i^m h_k^i \gamma_M^k) + \dots, \quad (\text{A12})$$

where the other terms indicated by the ellipsis indicate signed permutations of the free indices m, n, q, r and M, N, Q, R . Both terms can be evaluated in $\mathcal{O}(N^4)$ time by first defining the intermediate quantities

$$\mathcal{E}_1 = h_k^i \gamma_i^k, \quad (\text{A13a})$$

$$(\gamma_1 h \gamma_1)_M^m = \gamma_i^m h_k^i \gamma_M^k \quad (\text{A13b})$$

which can be evaluated cheaply though we will require the latter intermediate for all spinorbitals m and M and not just spinorbitals corresponding to the same spatial orbital.

Similar manipulations hold for the second term of Eq. (A10), $\gamma_{MNQRij}^{mnqrkl} v_{kl}^{ij}$. When expanding out the six-particle density matrix, we have three types of terms:

1. Dummy indices i and j could be on the same pair of one-particle density matrices as indices k and l , so that we have, for example, $\gamma_i^k \gamma_j^l \gamma_{MNQR}^{mnqr}$;

2. Only one of the dummy upper indices could be on the same one-particle density matrix as one of the dummy lower indices, so that we have terms such as $\gamma_M^k \gamma_j^l \gamma_i^m \gamma_{NQR}^{nqr}$;
3. Neither of the dummy upper indices could be on the same one-particle density matrix as one of the dummy lower indices, so that we have terms such as $\gamma_M^k \gamma_N^l \gamma_i^m \gamma_j^n \gamma_{QR}^{qr}$.

Respectively, these lead to contributions to $\bar{H}_{mn,qr}(\Omega)$ of the form

$$\delta \bar{H}_{mn,qr}(\Omega) \sim \gamma_{MNQR}^{mnqr} (v_{kl}^{ij} \gamma_i^k \gamma_j^l) + \dots, \quad (\text{A14a})$$

$$\delta \bar{H}_{mn,qr}(\Omega) \sim \gamma_{NQR}^{nqr} (\gamma_i^m \gamma_M^k \gamma_j^l v_{kl}^{ij}) + \dots, \quad (\text{A14b})$$

$$\delta \bar{H}_{mn,qr}(\Omega) \sim \gamma_{QR}^{qr} (\gamma_i^m \gamma_j^n \gamma_M^k \gamma_N^l v_{kl}^{ij}) + \dots, \quad (\text{A14c})$$

all of which can again be evaluated in $\mathcal{O}(N^4)$ time in terms of the intermediates

$$\mathcal{E}_2 = v_{kl}^{ij} \gamma_i^k \gamma_j^l, \quad (\text{A15a})$$

$$(\gamma_1 G \gamma_1)_M^m = \gamma_i^m \gamma_M^k \gamma_j^l v_{kl}^{ij}, \quad (\text{A15b})$$

$$(\gamma_2 v \gamma_2)_{MN}^{mn} = \gamma_i^m \gamma_j^n \gamma_M^k \gamma_N^l v_{kl}^{ij}. \quad (\text{A15c})$$

The last intermediate is itself computed in $\mathcal{O}(N^5)$ time. Similar manipulations follow for every other term in defining the Hamiltonian and overlap matrices.

Recall that there is an overall prefactor of N_{grid} arising from integrating $\bar{H}_{mn,qr}(\Omega)$, so the scaling is actually $\mathcal{O}(N^5 N_{\text{grid}})$. In the case of SUHF, the grid scales mildly with system size, but as mentioned before the integration grid in the case of lattice momentum projection itself scales with the number of sites, so that $N_{\text{grid}} \sim N$ and the scaling of JLM-SUHF is thus $\mathcal{O}(N^6 N_{\text{grid}}^{\text{SUHF}})$ where $N_{\text{grid}}^{\text{SUHF}}$ is the size of the grid needed for spin projection.

¹B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, *Chem. Phys.* **48**, 157 (1980).

²B. O. Roos, K. Andersson, M. P. Fulscher, P. A. Malmqvist, L. Serrana-Andres, K. Pierloot, and M. Merchan, *Adv. Chem. Phys.* **93**, 219 (1996).

³G. K.-L. Chan and S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 (2011).

⁴C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, *J. Chem. Phys.* **136**, 164109 (2012).

⁵P.-O. Löwdin, *Phys. Rev.* **97**, 1509 (1955).

⁶P. Ring and P. Schuck, *The Nuclear Many-Body Problem* (Springer-Verlag, New York, NY, 1980).

⁷J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems* (The MIT Press, Cambridge, MA, 1985).

⁸K. Schmid, *Prog. Part. Nucl. Phys.* **52**, 565 (2004).

⁹In general, we use not a projection operator *per se* but rather, something which acts as if it were. See Ref. 4 for details.

¹⁰R. Rodríguez-Guzmán, K. W. Schmid, C. A. Jiménez-Hoyos, and G. E. Scuseria, *Phys. Rev. B* **85**, 245130 (2012).

¹¹R. R. Rodríguez-Guzmán, C. A. Jiménez-Hoyos, R. Schutski, and G. E. Scuseria, *Phys. Rev. B* **87**, 235129 (2013).

- ¹²N. Tomita, *Phys. Rev. B* **69**, 045110 (2004).
- ¹³O. Juillet and R. Frésard, *Phys. Rev. B* **87**, 115136 (2013).
- ¹⁴S. Hirata, *Theor. Chem. Acc.* **129**, 727 (2011).
- ¹⁵E. Neuscamman, *Phys. Rev. Lett.* **109**, 203001 (2012).
- ¹⁶E. Neuscamman, “The Jastrow antisymmetric geminal power in Hilbert space: Theory, benchmarking, and application to a novel transition state,” preprint [arXiv:1308.5297](#) [cond-mat.str-el] (2013).
- ¹⁷E. Neuscamman, “A Jastrow factor coupled cluster theory for weak and strong electron correlation,” preprint [arXiv:1309.4137](#) [cond-mat.str-el] (2013).
- ¹⁸G. E. Scuseria, C. A. Jiménez-Hoyos, T. M. Henderson, J. K. Ellis, and K. Samanta, *J. Chem. Phys.* **135**, 124108 (2011).
- ¹⁹J. Hubbard, *Proc. R. Soc. London* **276**, 238 (1963).
- ²⁰T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).